DOCUMENT RESUME

ED 261 068 TM 850 463

AUTHOR Leonard, Tom; Novick, Melvin R.
TITLE Bayesian Inference and Diagnostics for the Three Parameter Logistic Model.
INSTITUTION Iowa Univ., Iowa City. Cada Research Group.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO ONR-TR-85-5
PUB DATE 8 Apr 85
CONTRACT N00014-83-C-0514
NOTE 30p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Bayesian Statistics; Cognitive Measurement; Estimation (Mathematics); *Latent Trait Theory; *Mathematical Models; Statistical Studies; *Test Theory
IDENTIFIERS *Birnbaum Models; *Three Parameter Model

ABSTRACT

This proposal attempts to follow in Allan Birnbaum's tradition by using Bayesian ideas to show that his mental test model possesses even broader applicability than previously realized. Birnbaum's two significant contributions to the theories of statistics and educational testing are: (1) the proof that the sufficiency and conditionality principles together imply the liklihood principle, thus justifying conditional or Bayesian inference when compared with frequentist inference; and (2) the mental test model introduced by Lord and Novick. A conditional maximization procedure is used here to approximate numerical integrations, thus broadening the scope of Bayesian inference and model checking. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

BEST COPY AVAILABLE

ED261068

BAYESIAN INFERENCE AND DIAGNOSTICS FOR THE THREE PARAMETER LOGISTIC MODEL

Tom Leonard
and
Melvin R. Novick

ONR Technical Report 85-5

CADA RESEARCH GROUP

TM 850 463

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

THE UNIVERSITY OF IOWA

BAYESIAN INFERENCE AND DIAGNOSTICS FOR THE THREE PARAMETER LOGISTIC MODEL

Tom Leonard
and
Melvin R. Novick

ONR Technical Report 85-5

April 8, 1985

This research was prepared under the Office of Naval Research Contract No. N00014-83-C-0514, Melvin R. Novick, Principal Investigator, The University of Iowa, Iowa City, Iowa.

Approved for public release, distribution unlimited. Reproduction in whole or part is permitted for any purpose of the United States Government.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY unclassified	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report 85-5	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Melvin R. Novick ✓	6b. OFFICE SYMBOL (If applicable)
6c. ADDRESS (City, State, and ZIP Code) 356 Lindquist Center The University of Iowa Iowa City, IA 52242 ✓	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research
	7b. ADDRESS (City, State, and ZIP Code) 536 South Clark Street Chicago, IL 60605
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Personnel & Training Res. Prog.	8b. OFFICE SYMBOL (If applicable)
	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-83-C-0514
8c. ADDRESS (City, State, and ZIP Code) Office of Naval Research (Code 458) Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	NR150-521		

11. TITLE (Include Security Classification)
Bayesian Inference and Diagnostics for the Three Parameter Logistic Model

12. PERSONAL AUTHOR(S) Tom Leonard and Melvin R. Novick

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM 7/1/83 TO 4/8/84	1985 April 8	27

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Bayesian inference logistic model exchangeability

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Allan Birnbaum made two historical contributions to the theories of statistics and educational testing. The first was his famous 1962 proof that the sufficiency and conditionality principles together imply the likelihood principle, thus justifying conditional, e.g., Bayesian, inference when compared with frequentist inference. The second was his famous mental test model introduced in Lord and Novick (1968). This proposal attempts to follow in Birnbaum's tradition by using Bayesian ideas to show that his mental test model possesses even broader applicability than hitherto realized. A conditional maximization procedure is used to approximate numerical integrations thus broadening the scope of Bayesian inference and model checking.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Charles Davis	22b. TELEPHONE (Include Area Code) (202) 696-4046 / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
unclassified

BAYESIAN INFERENCE AND DIAGNOSTICS FOR THE THREE PARAMETER LOGISTIC MODEL*

Tom Leonard
University of Wisconsin-Madison
and The University of Iowa
and
Melvin R. Novick
The University of Iowa

Abstract

Allan Birnbaum made two historical contributions to the theories of statistics and educational testing. The first was his famous 1962 proof that the sufficiency and conditionality principles together imply the likelihood principle, thus justifying conditional, e.g., Bayesian, inference when compared with frequentist inference. The second was his famous mental test model introduced in Lord and Novick (1968). This proposal attempts to follow in Birnbaum's tradition by using Bayesian ideas to show that his mental test model possesses even broader applicability than hitherto realized. A conditional maximization procedure is used to approximate numerical integrations thus broadening the scope of Bayesian inference and model checking.

*Support for this research was provided under contract #N00014-83-C-0514 with the Personnel Training Branch of the Office of Naval Research.

Bayesian Inference and Diagnostics

For The Three Parameter Logistic Model

1. Birnbaum's Model

For $i = 1, \ldots, r$ and $j = 1, \ldots, s$, let y_{ij} represent the binary response of examinee i on item j, with

$$p(y_{ij} = 1) = \theta_{ij} \; ; \; p(y_{ij} = 0) = 1 - \theta_{ij} \; .$$

Birnbaum (1969) considers the model

$$\gamma_{ij} = \log \theta_{ij} - \log(1 - \theta_{ij}) = a_j \, (\alpha_i - \beta_j) \qquad (1.1)$$

$$(i = 1, \ldots, r; \; j = 1, \ldots, s)$$

where α_i is the ability of examinee i, β_j is the difficulty of item j, and a_j is the discrimination parameter for item j.

Rather than estimating the x's, β's, and a's by maximum likelihood (Lord and Novick, 1968), it is possible to refer to a Bayesian procedure which incorporates prior information with the information in the data. Leonard (1973a) considers a multiplicative Poisson model due to Rasch (1960). Birnbaum (1969), Owen (1975), and Swaminathan and Gifford (1980, 1981, 1982) propose some interesting Bayesian methods for Birnbaum's and related models.

For example, Swaminathan and Gifford (1981) construct a normal/normal/Gamma/beta prior assuming exchangeability of the α_i, of the β_j, and of the a_j, and are able to compute point estimates based on the joint posterior modes. They also include the important generalization

$$\theta_{ij} = g_j + (1 - g_j) \frac{e^{a_j(\alpha_i - \beta_j)}}{1 + e^{a_j(\alpha_i - \beta_j)}} \qquad (1.2)$$

where g_j is the guessing parameter for the gth item. There are some general problems with joint posterior modes in Beyesian inference. See O'Hagan (1976).

Birnbaum (1969) assumes a logistic prior distribution for ability and Owen (1978) considers Bayesian sequential updating for the two parameter normal ogive model.

The approaches outlined in this proposal have the following potential advantages when compared with existing methodology in the literature.

(A) A much more flexible family of prior distributions is available than hitherto appreciated in the literature using the ideas initiated by Leonard (1972, 1973) and Novick, Lewis and Jackson (1973) of transforming to new

parameters which may be taken to possess a multivariate normal distribution. Exchangeable priors are included as special cases, but many prior structures are available, e.g., for the discrimination and guessing parameters. It is also possible to proceed under prior uncertainty in which case (B) extends current maximum likelihood methodology.

(B) Using ideas considered by Leonard (1982) and Tierney and Kadane (1984) it is possible to compute close approximations to the entire marginal posterior distributions of the parameters of interest, using a conditional maximization procedure. This avoids numerical integrations even when the number of parameters is large. It is therefore possible to draw general inferences about the parameters e.g., using exchangeable or more general inferences prior distributions rather than just relying on point estimates. Related methodology is described by Leonard and Novick (1984) in a two-way contingency table situation. As well as drawing general inferences it is possible, if required, to refer to full decision theoretic methods; this is difficult under existing techniques for point estimation.

(C) Following the philosophy of Leonard and Novick (1984) it is possible to consider a generalization of Birnbaum's model permitting the statistician to express uncertainty about the original model.(See section 4) By considering the posterior distribution of certain "parametric residuals" it is then possible to judge whether Birnbaum's model provides a reasonable fit in the data, and to isolate items which deviate from Birnbaum's model. Similar procedures can be used to judge modifications or alternatives to Birnbaum's model, e.g., (1.2).

(D) In the special case where the α's, β's, and a's are separately exchangeable it is possible to relax the assumption that their prior distribution belongs to a specialized family and to use a hierarchical procedure which effectively estimates these mixing distributions from the data, following the approach, for a single set of exchangeable parameters,

by Leonard (1984). This approach will be particularly valuable when either the examinees, or the items, are distinguished to be in two or more groups, or when one item yields different performances than the remainder. The posterior estimates do not inevitably shrink towards a common value in standard Bayesian fashion, but instead take more account of the scatter and grouping of the data, and provide a procedure close to cluster analysis.

Existing Computer Packages

The two most frequently applied computer programs for this model are (1) LOGIST which is based upon the broad methodology of Lord (1980) e.g., it efficiently computes joint maximum likelihood estimates of the parameters, and (2) BILOG which uses the EM algorithm to compute the marginal maximum likelihood estimates of the item parameters under an assumed mixing distribution for the ability parameters - see Darrell Bock and Aitken (1981).

Our own methodology can be regarded as in the spirit of the EM algorithm since it employs a conditional maximization to approximate a marginalization. The key difference is that we are able to approximate complete marginal posterior densities and hence perform a general inference rather than just using point estimates and standard deviations. This yields a wide range of measures of information about individual parameters of interest.

Darrell Bock (1972) proposes pioneering work when responses are found in two or more nominal categories. Our procedures could, if required, be extended to his multivariate logit models - see the method by Leonard (1973) for smoothing the probabilities in a histogram.

2. The Prior Distribution

Swaminathan and Gifford assume normal priors for the α_i and β_j, a chi-square prior for the a_j and a beta prior for the g_j, if guessing parameters are incorporated. We instead recommend the transformations

$$\xi_j = \log a_j, \quad \lambda_j = \log g_j - \log(1 - g_j) \tag{2.1}$$

and taking the vectors $\underset{\sim}{\alpha} = (\alpha_1, \ldots, \alpha_r)^T$, $\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_s)^T$, $\underset{\sim}{\xi} = (\xi_1, \ldots, \xi_s)^T$, and $\underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_s)^T$ to possess independent multivariate normal distributions, conditional on their respective mean vectors $\underset{\sim}{\mu}_\alpha$, $\underset{\sim}{\mu}_\beta$, $\underset{\sim}{\mu}_\xi$, $\underset{\sim}{\mu}_\lambda$ and covariance matrices $\underset{\sim}{\Sigma}_\alpha$, $\underset{\sim}{\Sigma}_\beta$, $\underset{\sim}{\Sigma}_\xi$, and $\underset{\sim}{\Sigma}_\lambda$.

Owing to the large number of prior parameters appearing in the mean vectors and covariance matrices this prior specification is much more general than previous suggestions in the literature and will permit a variety of interesting special cases. Moreover, a satisfactory posterior analysis is readily computable. The joint posterior density of $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\xi}$, and $\underset{\sim}{\lambda}$ is given by

$$\pi (\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\xi}, \underset{\sim}{\lambda} \mid \underset{\sim}{\chi}, \underset{\sim}{\mu}, \underset{\sim}{C})$$

$$\propto \frac{\prod_{ij} \left[e^{\lambda_j} + (1 + e^{\lambda_j}) \; e^{a_j(\alpha_i - \beta_j)} \right]^{y_{ij}}}{\prod_{ij} (1 + e^{a_j(\alpha_i - \beta_j)}) \; \prod_{j} (1 + e^{\lambda_j})^r}$$

$$\times \exp \left[-\tfrac{1}{2} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}_{\alpha}) \, \underset{\sim}{C}_{\alpha}^{-1} \, (\underset{\sim}{\alpha} - \underset{\sim}{\mu}_{\alpha}) - \tfrac{1}{2} (\underset{\sim}{\beta} - \underset{\sim}{\mu}_{\beta}) \, \underset{\sim}{C}_{\beta}^{-1} \, (\underset{\sim}{\beta} - \underset{\sim}{\mu}_{\beta}) \right. \tag{2.2}$$

$$- \tfrac{1}{2} (\underset{\sim}{\xi} - \underset{\sim}{\mu}_{\xi}) \, \underset{\sim}{C}_{\xi}^{-1} \, (\underset{\sim}{\xi} - \underset{\sim}{\mu}_{\xi})$$

$$\left. - \tfrac{1}{2} (\underset{\sim}{\lambda} - \underset{\sim}{\mu}_{\lambda}) \, \underset{\sim}{C}_{\lambda}^{-1} \, (\underset{\sim}{\lambda} - \underset{\sim}{\mu}_{\lambda}) \right]$$

and it is possible to compute close approximations to all the marginal densities, using the ideas outlined in section 3.

When the four different sets of parameters are exchangeable, we may take $\underset{\sim}{\mu}_{\alpha} = 0$, $\underset{\sim}{\mu}_{\beta} = \mu_{\beta} \underset{\sim}{e}_s$, $\underset{\sim}{\mu}_{\xi} = \mu_{\xi} \underset{\sim}{e}_s$, $\underset{\sim}{\mu}_{\lambda} = \mu_{\lambda} \underset{\sim}{e}_s$, $\underset{\sim}{C}_{\alpha} = \sigma_{\alpha}^2 \, \underset{\sim}{I}_r$, $\underset{\sim}{C}_{\beta} = \sigma_{\beta}^2 \, \underset{\sim}{I}_s$, $\underset{\sim}{C}_{\xi} = \sigma_{\xi}^2 \, \underset{\sim}{I}_s$, and $\underset{\sim}{C}_{\lambda} = \sigma_{\lambda}^2 \underset{\sim}{I}_s$ where μ_{β}, μ_{ξ}, and μ_{λ} are scalars and $\underset{\sim}{e}_s$ is an s x 1 vectors of ones. A hierarchical Bayesian approach is now available where, say, ignorance priors are assigned to the prior parameters μ_{β}, μ_{ξ}, μ_{λ}, σ_{α}^2, σ_{β}^2, σ_{ξ}^2, and σ_{λ}^2, so that the posterior evaluation of these quantities is based largely on the data. This will provide Stein-type shrinkage estimates for the α_i, the β_j, the ξ_j, and the λ_j, smoothing each of these four sets of parameters towards

an average value. From experience with shrinkage estimates in other situations, it is evident that they will possess vastly superior mean squared error properties when compared with maximum likelihood.

A variety of other flexible prior structures are available. For example, if the items are to be attempted in order it might be reasonable to take C_{β}, C_{ξ}, and C_{λ} to possess elements C_{β}^{ij}, C_{ξ}^{ij}, and C_{λ}^{ij}, satisfying

$$C_{\beta}^{ij} = \sigma_{\beta}^2 \rho_{\beta}^{|i-j|}$$

$$C_{\xi}^{ij} = \sigma_{\xi}^2 \rho_{\xi}^{|i-j|}$$

$$C_{\lambda}^{ij} = \sigma_{\lambda}^2 \rho_{\lambda}^{|i-j|} \quad (i, j = 1, \ldots, s)$$

i.e. possessing autoregressive correlation structures in the manner recommended by Leonard (1973) for smoothing the cell probabilities in a histogram.

Alternatively, we might have hypothesized regression models for α, β, ξ, and λ which we represent by taking their prior means to satisfy $\mu_{\alpha} = X_{\alpha}\zeta_{\alpha}$, $\mu_{\beta} = X_{\beta}\zeta_{\beta}$, $\mu_{\xi} = X_{\zeta}\zeta_{\zeta}$, and $\mu_{\lambda} = X_{\lambda}\zeta_{\lambda}$ where the X's are specified design vectors and the ζ's are unknown vectors of regression coefficients. We may represent uncertainty in these regressions by perhaps taking $C_{\alpha} = \sigma_{\alpha}^2 I_{\alpha}$, $C_{\beta} = \sigma_{\beta}^2 I_{\beta}$, $C_{\xi} = \sigma_{\xi}^2 I_{\xi}$, and $C_{\lambda} = \sigma_{\lambda}^2 I_{\lambda}$ and then assigning ignorance priors to the ζ's and σ's.

It seems difficult to incorporate regressors for the discrimination or guessing parameters using any of the previous Bayesian approaches in the literature.

3. Marginalization Procedures

We will be able to handle the multiparameter Bayesian models introduced in section 2 because of (a) the large increase in computer efficiency over the last few years and (b) a technical device enabling us to replace numerical integrations by a conditional maximization. This is simply the first item in an Edgeworth expansion but its impact upon Bayesian marginalization does not seem to have been noticed until Leonard (1982).

Suppose, in general, that an m x 1 vector of parameters $\underset{\sim}{\alpha} = (\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2)$ possesses posterior density $\pi(\underset{\sim}{\alpha}|\underset{\sim}{y})$ where $\underset{\sim}{\alpha}_1$ is a p x 1 subvector of interest and $\underset{\sim}{\alpha}_2$ is the q x 1 subvector of the remaining q x 1 nuisance parameters. Then, under fairly wide regularity conditions the marginal posterior density of $\underset{\sim}{\alpha}_1$ is closely approximated by

$$\pi(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2 = \hat{\underset{\sim}{\alpha}}_2(\underset{\sim}{\alpha}_1) | \underset{\sim}{y}) \;/\; R(\underset{\sim}{\alpha}_1) \tag{3.1}$$

where $\hat{\underset{\sim}{\alpha}}_2(\underset{\sim}{\alpha}_1)$ is the conditional posterior mode of $\underset{\sim}{\alpha}_2$ given $\underset{\sim}{\alpha}_1$, and

$$R(\underset{\sim}{\alpha}_1) = (2\pi)^{-\frac{1}{2}p} \left| \frac{\partial^2 \log \pi(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2 \mid \underset{\sim}{y})}{\partial(\underset{\sim}{\alpha}_2 \underset{\sim}{\alpha}_2)^T} \right|^{\frac{1}{2}}_{\underset{\sim}{\alpha}_2 = \hat{\underset{\sim}{\alpha}}_2(\underset{\sim}{\alpha}_1)} \tag{3.2}$$

Therefore, for each value of $\underset{\sim}{\alpha}_1$ (typically p = 1 with $\underset{\sim}{\alpha}_1$ a scaler) instead of performing a q-dimensional numerical integration with respect to $\underset{\sim}{\alpha}_2$ we just conditionally maximize (e.g., using Newton-Raphson) the joint posterior density with respect to $\underset{\sim}{\alpha}_2$, and then adjust the supremum by the relatively minor determinant (of the Hessian) term in the denominator. This procedure can be shown to provide surprisingly accurate results in virtually any situation where the integrations can also be performed analytically.

It seems possible to apply this procedure to the situation introduced in section 2, e.g., to the joint posterior density in (2.2) in order to investigate the marginal posterior density of any specific item or examinee parameter. In hierarchical Bayes situations where the estimation of prior parameters is involved it is anticipated that it will be superior to combine the above conditional maximizations with some low dimensional numerical integrations with respect to the important prior parameters. Improvements and approximations to the basic conditional maximization procedure are envisioned. Similar approximations to posterior moments and Bayes estimates are available.

4. A Generalization of Birnbaum's Model

For the purpose of checking the adequacy of Birnbaum's model we consider the generalization which replaces (1.1) by the (sampling) assumption

$$\gamma_{ij} = \log \theta_{ij} - \log (1-\theta_{ij}) = a_j (\alpha_i - \beta_j) + \varepsilon_{ij} \qquad (4.1)$$

$$(i=1, \ldots, r;\ j=1, \ldots, s)$$

where the ε_{ij} are independent error terms which are normally distributed with zero means and common variance σ_γ^2. A positive value for σ_γ^2 permits the statistician to express uncertainty as to whether Birnbaum's model is appropriate. A data based estimate for σ_γ^2 helps the statistician to judge whether Birnbaum's model provides a good fit to the data. In the spirit of Leonard and Novick (1984), the quantity $\varepsilon_{ij} = \gamma_{ij} - a_j(\alpha_i - \beta_j)$ could be called a "parametric residual". Posterior inferences about the ε_{ij} permit the statistician to judge specific deviations from Birnbaum's model.

The joint likelihood of the parameters of interest is now

$$\ell\ (\underset{\sim}{a},\ \underset{\sim}{\alpha},\ \underset{\sim}{\beta},\ \sigma_\gamma^2,\ \underset{\sim}{\gamma} \mid \underset{\sim}{y})$$

$$\propto (\sigma_\gamma^2)^{-\frac{1}{2}rs} \exp\left\{ \sum_{ij} y_{ij}\gamma_{ij} - \tfrac{1}{2}\sigma_\gamma^2 \sum_{ij} (\gamma_{ij} - a_j(\alpha_i - \beta_j))^2 \right\} \Big/ \prod_{ij} (1 + e^{\gamma_{ij}}) \qquad (4.2)$$

and, with appropriate marginalizations, it is possible to draw inferences e.g., about σ_γ^2 and the ε_{ij}, by combining this likelihood with any prior information about the a's, α's, β's, and σ_γ^2's, using the prior structures introduced in section 2.

As a special case, suppose we have uniform priors for $\log \sigma_\gamma^2$, the $\xi_j = \log a_j$, the α_i, and the β_j. Then a set of posterior modes of the ξ_j, α_i, β_j, and γ_{ij}, given σ_γ^2, satisfies

$$e^{\tilde{\gamma}_{ij}} = 1 - \sigma_\gamma^{-2} \left(\tilde{\gamma}_{ij} - e^{\tilde{\xi}_j} \left(\tilde{\alpha}_i - \tilde{\beta}_j\right)\right) \qquad (i = 1, \ldots, r;\ j = 1, \ldots, s) \tag{4.3}$$

$$\tilde{\alpha}_i = \sum_j e^{\tilde{\xi}_j} \left(\tilde{\gamma}_{ij} - \tilde{\gamma}_{.j}\right) \Big/ \sum_j e^{\tilde{\xi}_j} \qquad (i = 1, \ldots, r) \tag{4.4}$$

$$\tilde{\beta}_j = -\tilde{\gamma}_{.j} / e^{\tilde{\xi}_j} \qquad (j = 1, \ldots, s) \tag{4.5}$$

$$e^{\tilde{\xi}_j} = \sum_i \tilde{\gamma}_{ij} \left(\tilde{\alpha}_i - \tilde{\beta}_j\right) / \sum_i \left(\tilde{\alpha}_i - \tilde{\beta}_j\right)^2 \qquad (j = 1, \ldots, s) \tag{4.6}$$

which we readily solved computationally. The marginal posterior density of σ_γ^2 may hence be approximated via the approach described in section 3, providing a useful inferential procedure for checking Birnbaum's model ($\sigma_\gamma^2 = 0$).

It is hoped to be able to combine this procedure with the hierarchical prior discussed in section 2 for the parameters within Birnbaum's model. The examinee and item parameter estimates are open to smoothing both by an expression of disbelief in Birnbaum's model, and an expression of, say, exchangeability, about the parameters themselves.

The ideas of this section may be used to also generalize the four parameter model in (1.2) which may be expressed in the form

$$\gamma_{ij} = \log \left[e^{\lambda_j} + (1 + e^{\lambda_j})\, e^{e^{\xi_j}(\alpha_i - \beta_j)} \right] \tag{4.7}$$

where $\gamma_{ij} = \log \theta_{ij} - \log (1-\theta_{ij})$ and λ_j satisfies (2.1).

We simply replace (4.7) by the assumption

$$\gamma_{ij} = \gamma_{ij}^{*} + \varepsilon_{ij} \tag{4.8}$$

whose γ_{ij}^{*} denotes the expression in (4.7) and

where the ϵ_{ij} are the same independent error terms as before. A full Bayesian analysis of the extended model in (4.8), with fourfold exchangeability (on the α's, β's, ξ's, and λ's) is one of the mean objectives of our research. It will be necessary to estimate the variance σ_γ^2 together with the eight prior parameters μ_α, μ_β, μ_ξ, μ_λ, σ_α^2, σ_β^2, σ_ξ^2, and σ_λ^2, using the marginalization procedures previously discussed.

5. Relaxing Prior Normality

Suppose, in general, that parameters $\alpha_1, \ldots, \alpha_r$ are a priori exchangeable, and taken to be independent, given a density $g(\alpha)$, and to each possess this density. Frequently (e.g., Swaminathan and Gifford, Leonard 1973) this distribution is either taken to be normal, or conjugate, or some transformation of normal. It is typically assumed to be unimodal with thin tails. However, the functional form assumed for g has a pronounced effect upon the posterior estimates of $\alpha_1, \ldots, \alpha_r$. Thicker tails would cause outlying observations to be pulled in less. Bimodality tends to cause parameters to be smoothed within two separate groups. Discrete distributions can cause the parameters to be isolated into clusters.

Laird (1978) and Leonard (1984) go to the other extreme and estimate the density g entirely from the data. If observations $y_1, \ldots, y_r$ are independent, given $\alpha_1, \ldots, \alpha_r$, with respective densities $p(y_1 \mid \alpha_1), \ldots, p(y_r \mid \alpha_r)$, then g can be estimated by maximizing its marginal likelihood

$$\ell\,(g \mid \chi) = \prod_{i=1}^{r} \left[\int p(y_i|\alpha)\; g(\alpha)\; d\alpha \right] \qquad (5.1)$$

The maximum likelihood estimate of g is a discrete distribution, and relatively easy to compute. In a series of numerical examples in the binomial exchangeability situation Leonard (1984) shows that the consequent (empirical) Bayesian estimates of the parameters of interest can both differ numerically from more standard Bayes shrinkage estimates, and take useful account of the scatter of the data. The clustering effect of the posterior estimates seems particularly useful for Birnbaum's model, or the generalization in section 4, since the examinees may be grouped into clusters or the difficulty of one or two items may turn out to be extreme.

It is intended to permit g to be discrete and then to construct a suitable prior distribution for g (see Leonard, 1978) thus providing a more general Bayesian alternative to Leonard's previous empirical Bayesian approach. It will be interesting to apply these items to the examinee and item parameters in Birnbaum's model, i.e. to three sets of exchangeable parameters simultaneously.

Conclusions

After fifteen years several useful Bayesian attempts have been made to analyze Birnbaum's model. However, with rapid new developments in Bayesian methodology there is still considerable scope for a more thorough treatment of this model. The objective is to provide procedures for the analysis of data which are not available via the non-Bayesian medium

and to provide improvements on currently available Bayesian methodologies. In particular we suggest that this line of research will provide better and more flexible prior distributions. improved computational algorithms, methods of model checking, and improved decision inference procedures for the selection of items.

In the proposed approach most subjectivity will be incorporated in the hierarchical specification of meaningful prior structures i.e. the functional forms of multi-parameter prior distributions. Prior parameters will frequently be evaluated with the assistance of the data by assigning them ignorance priors at a lower stage of the prior model. Hence the user should find the procedures straight forward to implement when analyzing his data.

A key practical application (e.g., Owen 1975) is the development of methods for choosing an examinee's next item from a choice of possible items. One criterion is to choose the item for which the predicted probability of success is closest to ½. Our approach tackles this problem in a straight-forward manner. Simply average the probability in (1.2) with respect to the marginal posterior density of the ability parameter, and, for each item, with respect to the posterior density of the difficulty, discrimination, and test parameters. Then compare with either ½ or the predicted value of $g_j + \frac{1}{2}(1 - g_j)$.

It would alternatively be possible to use our marginalization procedures to approximate the full posterior density of the probability in (1.2) for each potential item separately. This would permit general inferences about potential items and present a practical alternative to

Owen's decision theoretic procedure which involves assumptions regarding expected loss. Full decision theoretic procedures, generalizing Owen's, are also available.

REFERENCES

Birnbaum, A. (1969) Statistical theory for logistic mental test models with a prior distribution of ability. J. Math. Psych, 6, 258-276.

Darrell Bock, R. (1972) Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 37, 29-51.

Darrell Bock, R. & Aitken, M. (1981) Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. Psychometrika, 45, 443-460.

Laird, N. M. (1978) Non parametric estimation of a mixing distribution. J. Amer. Stat. Assoc., 73, 805-811.

Leonard, T. (1982) Bayesian methods for binomial data, Biometrika, 59, 581-589.

Leonard, T. (1973a) Bayesian methods for the simultaneous estimation of several parameters, Ph.D. thesis. University College, London.

Leonard, T. (1973b) A Bayesian method for histograms, Biometrika, 60, 297-308.

Leonard, T. (1978) Density estimation, stochastic processes, and prior information (with Discussion) J. Roy. Stat. Soc. B.

Leonard, T. (1982) Comment on the paper by Lejeune and Faulkenberry, J. Amer. Stat. Assoc.

Leonard, T. (1984) Some data analytic modifications to Bayes-Stein estimation. Annals of Inst. Stat. Math. (to appear).

Leonard, T. and Novick, M. R. (1984) Bayesian marginalization for two-way contingency tables. Lindquist Center for Measurement, The University of Iowa, Technical Report (to appear).

Lord, F. M. (1980) Applications of item response theory to practical testing problems. LEA.

Lord, F. M. & Novick, M. R. (1968) Statistical theories of mental test scores. Addison-Wesley.

O'Hagan, A. (1970) On posterior joint and marginal modes. Biometrika, 63, 329-334.

Owen, R. (1975) A Bayesian sequential procedure for quantal response in the context of adaptive testing. J. Amer. Stat. Assoc. 70, 351-356.

Swaminathan, H. & Gifford, J. A. (1980) Bayesian estimation in the one parameter latent trait model, Laboratory of Psychometric Research, Report No. 106, University of Massachusetts.

Swaminathan, H. & Gifford, J. A. (1981) Bayesian estimation in the three parameter logistic model. Laboratory of Psychometric Research, Report No. 119, University of Massachusetts.

Swaminathan, H. & Gifford, J. A. (1982) Bayesian estimation in item response models. School of Education Research, Report LR-120, University of Massachusetts.

Tierney, L. and Kadane, J. (1984) Approximating marginal posterior moments and densities. Paper presented to seminar on Bayesian Inference in Econometrics, Madison, Wisconsin, May, 1984.

Distribution List

Personnel Analysis Division
AF/MPXA
5C360, The Pentagon
Washington, DC 20330

Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

Air Force Office
of Scientific Research
Life Sciences Directorate
Bolling Air Force Base
Washington, DC 20332

Dr. James Algina
University of Florida
Gainesville, FL 32605

Technical Director
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Special Assistant for Projects
OASN(M&RA)
5D800, The Pentagon
Washington, DC 20350

Dr. R. Darrell Bock
University of Chicago
Department of Education
Chicago, IL 60637

Dr. Nick Bond
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

Dr. Robert Brennan
American College Testing
Programs
P. O. Box 168
Iowa City, IA 52243

Col. Roger Campbell
AF/MPXOA
Pentagon, Room 4E195
Washington, DC 20330

Dr. James Carlson
American College Testing
Program
P.O. Box 168
Iowa City, IA 52243

Mr. Raymond E. Christal
AFHRL/MOE
Brooks AFB, TX 78235

Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

Director
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Scientific Advisor
to the DCNO (MPT)
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Chief of Naval Education
and Training
Liason Office
AFHRL
Operations Training Division
Williams AFB, AZ 85224

Office of the Chief
of Naval Operations
Research Development
& Studies Branch
NAVOP 0'B7
Washington, DC 20350

Dr. Stanley Collyer
Office of Naval Technology
800 N. Quincy Street
Arlington, VA 22217

CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217

Mr. Timothy Davey
University of Illinois
Educational Psychology
Urbana, IL 61801

Dr. Dattprasad Divgi
Syracuse University
Department of Psychology
Syracuse, NY 13210

Dr. Fritz Drasgow
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Stephen Dunbar
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

Dr. John M. Eddins
University of Illinois
252 Engineering Research
Laboratory
103 South Mathews Street
Urbana, IL 61801

Dr. Richard Elster
Deputy Assistant Secretary
of the Navy (Manpower)
Washington, DC 20350

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Dr. Benjamin A. Fairbank
Performance Metrics, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

Dr. Marshall J. Farr
2520 North Vernon Street
Arlington, VA 22207

Dr. Leonard Feldt
Lindquist Center
for Measurment
University of Iowa
Iowa City, IA 52242

Dr. Myron Fischl
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Alfred R. Fregly
AFOSR/NL
Bolling AFB, DC 20332

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 62
San Diego, CA 92152

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

Mr. Dick Hoshaw
NAVOP-135
Arlington Annex
Room 2834
Washington, DC 20350

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Dr. Huynh Huynh
College of Education
Univ. of South Carolina
Columbia, SC 29208

Dr. Douglas H. Jones
Advanced Statistical
Technologies Corporation
10 Trafalgar Court
Lawrenceville, NJ 08148

Dr. William Koch
University of Texas-Austin
Measurement and Evaluation
Center
Austin, TX 78703

Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Anita Lancaster
Accession Policy
OASD/MI&L/MP&FM/AP
Pentagon
Washington, DC 20301

Dr. Marcy Lansman
University of North Carolina
The L. L. Thurstone Lab.
Davie Hall 013A
Chapel Hill, NC 27514

Dr. Jerry Lehnus
OASD (M&RA)
Washington, DC 20301

Dr. Thomas Leonard
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53705

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
The NETHERLANDS

Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

Dr. Robert Lockman
Center for Naval Analysis
200 North Beauregard St.
Alexandria, VA 22311

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen Martin
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

Dr. James McBride
Psychological Corporation
c/o Harcourt, Brace,
Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. Clarence McCormick
HQ, MEPCOM
MEPCT-P
2500 Green Bay Road
North Chicago, IL 60064

Mr. Robert McKinley
University of Toledo
Dept of Educational Psychology
Toledo, OH 43606

Dr. Barbara Means
Human Resources
Research Organization
300 North Washington
Alexandria, VA 22314

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. Karen Mitchell
Army Research Institute
5001 Eisenhower Blvd
Alexandria, VA 22333

Ms. Kathleen Moreno
Navy Personnel R&D Center
Code 62
San Diego, CA 92152

Headquarters, Marine Corps
Code MPI-20
Washington, DC 20380

Lt. Col. Jim Murphy
HQ, Marine Corps
Code MRRP
Washington, DC 20380

Director
Research & Analysis Division
Navy Recruiting Command (Code 22)
4015 Wilson Blvd.
Arlington, VA 22203

Program Manager for Manpower,
Personnel, and Training
NAVMAT 0722
Arlington, VA 22217

Assistant for MPT Research,
Development and Studies
NAVOP 01B7
Washington, DC 20370

Head, Military Compensation
Policy Branch
NAVOP 134
Washington, DC 20370

Dr. W. Alan Nicewander
University of Oklahoma
Department of Psychology
Oklahoma City, OK 73069

Director, Manpower and Personnel
Laboratory
NPRDC (Code 06)
San Diego, CA 92152

Library
Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. James Olson
WICAT, Inc.
1875 South State Street
Orem, UT 84057

Group Psychology Program
Code 442GP
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217-5000

Office of Naval Research
Code 442PT
800 N. Quincy Street
Arlington, VA 22217
(6 Copies)

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

Commanding Officer
Army Research Institute
ATTN: PERI-BR (Dr. J. Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

Military Assistant for Training
and Personnel Technology
OUSD(R&E)
Room 3D129, The Pentagon
Washington, DC 20301

Dr. Randolph Park
AFHRL/MOAN
Brooks AFB, TX 78235

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Cirle, NW
Washington, DC 20036

Daira Paulson
Code 52 - Training Systems
Navy Personnel R&D Center
San Diego, CA 92152

Dr. James Paulson
Dept. of Psychology
Portland State Unive3rsity
P.O. Box 751
Portland, OR 97207

Dr. Sam Pearson
CACI, Inc. FED.
Department 5921
1815 North Fort Myers Dr.
Arlington, VA 22209

Dr. James W. Pellegrino
University of California,
Santa Barbara
Dept. of Psychology
Santa Barabara, CA 93106

Administrative Sciences Dept.
Naval Postgraduate School
Monterey, CA 93940

Department of Operations Research
Naval Postgraduate School
Monterey, CA 93940

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
Knoxville, TN 37916

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152

Dr. Robert Sasmor
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

Dr. Mary Schratz
Navy Personnel R&D Center
San Diego, CA 92152

Dr. W. Steve Sellman
OASD(MRA&L)
2B269 The Pentagon
Washington, DC 20301

Dr. Joyce Shields
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujusawa 251
JAPAN

Dr. William Sims
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Alfred F. Smode
Senior Scientist
Code 7B
Naval Training Equipment Center
Orlando, FL 32813

Dr. Richard Snow
Liaison Scientist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

Maj. Bill Strickland
AF/MPXOA
4E168 Pentagon
Washington, DC 20330

Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E. Street, NW
Washington, DC 20415

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 310
St. Paul, MN 55114

Dr. Frank Vicino
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Howard Wainer
Division of Psychological Studies
Educational Testing Service
Princeton, NJ 08540

Dr. Ming-Mei Wang
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Mr. Thomas A. Warm
Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Dr. Brian Waters
HumRRO
300 North Washington
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

Dr. Ronald Weitzman
Naval Postgraduate School
Administrative Sciences Dept.
Monterey, CA 93940

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90007

German Military Representative
ATTN: Wolfgang Wildegrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

Dr. Bruce Williams
Department of Educational
Psychology
University of Illinois
Urbana, IL 61801

Dr. Hilda Wing
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

Major Frank Yohannan, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380